

Stochastic simulation of precipitation-consistent daily reference evapotranspiration using vine copulas

M. T. Pham¹ · H. Vernieuwe²  · B. De Baets² · P. Willems³ · N. E. C. Verhoest¹

Published online: 24 November 2015

© The Author(s) 2015. This article is published with open access at Springerlink.com

Abstract Evapotranspiration is an important process in the water cycle that represents a considerable amount of moisture lost to the atmosphere through evaporation from the soil and wet surfaces, and transpiration from plants. Therefore, several water management methods, such as irrigation scheduling and hydrological impact analysis, rely on an accurate estimation of evapotranspiration rates. Often, daily reference evapotranspiration is modelled based on the Penman, Priestley–Taylor or Hargraeves equation. However, each of these models requires extensive input data, such as daily mean temperature, wind speed, relative humidity and solar radiation. Yet, in design studies, such data may be unavailable and therefore, another approach may be needed that is based on stochastically generated time series. More specifically, when rainfall-runoff models are used, these evapotranspiration data need to be consistent with the accompanying (stochastically generated) precipitation time series data. In this paper, such an approach is presented in which the statistical dependence between evapotranspiration, precipitation and temperature is described by three- and four-dimensional vine copulas.

Based on a case study of 72 years of evapotranspiration, temperature and precipitation data, observed in Uccle, Belgium, it is shown that canonical vine copulas (C-vines) perform very well in preserving the dependences between variables.

Keywords Evapotranspiration · Vine copula · Stochastic modelling · Hydrological impact analysis

1 Introduction

Water management is often concerned with preventing or mitigating extreme conditions. Therefore, hydraulic structures, such as flood control reservoirs, are often constructed on rivers in order to prevent floods or to mitigate their consequences. Alternative structures, such as water supply reservoirs or irrigation structures, should provide water in times of drought, during which the available amount of fresh water hardly meets the water demand. It is therefore of major importance to accurately dimension these structures such that they can cope with a hazard, being a flood or a drought event, of a given magnitude, duration and frequency of occurrence or its return period. In practice, this can be accomplished by the use of design storms with given statistical properties (Wheater 2002; Willems 2013) or the use of long-term rainfall records in order to obtain a continuous discharge series from a rainfall-runoff model, from which flood or low flow events are extracted (Verhoest et al. 2010; Willems 2014). The latter approach is preferred as it takes into account the antecedent wetness state of the catchment, and hence yields more reliable statistics. Yet, in order to cope with low frequency events, very long precipitation records are needed, which are generally not available in practice. This can be overcome

✉ H. Vernieuwe
hilde.vernieuwe@ugent.be

M. T. Pham
minhtu.pham@ugent.be

¹ Laboratory of Hydrology and Water Management, Ghent University, Coupure links 653, 9000 Ghent, Belgium

² KERMIT, Department of Mathematical Modelling, Statistics and Bioinformatics, Ghent University, Coupure links 653, 9000 Ghent, Belgium

³ Civil Engineering Department, KU Leuven, Kasteelpark Arenberg 40, Box 2448, 3001 Leuven, Belgium

by using a stochastically modelled rainfall time series (Boughton and Droop 2003). Furthermore, as the catchment discharge is also influenced by the amount of water that is evaporated, it is also important to employ a time series of evapotranspiration values.

From a physics point of view, evapotranspiration is determined by several climatological variables, including net radiation, wind speed, air temperature, air humidity and air pressure. Classical evapotranspiration models that use equations such as the Penman, Priestley–Taylor or Hargraeves equation, need extensive input of these variables to generate time series of evapotranspiration values. Furthermore, as all of these variables are stochastic by nature, evapotranspiration therefore also is a stochastic variable. However, as commented by Srikanthan and McMahon (2001), the stochastic modelling of climate data should preserve the cross-correlation or dependence between variables. In this sense, the correlation structure between evapotranspiration and precipitation should be maintained when generating both time series as input to a hydrological model. Jones et al. (1972) already hypothesized that daily evaporation is related to the day of the year and the precipitation of the day in question and the preceding day. At large time scales (yearly) evapotranspiration has been shown to be related to precipitation, as expressed by the Budyko curve (Arora 2002; Gerrits et al. 2009). However, at the daily time scale, this correlation is generally not explicitly taken into account for modelling evapotranspiration. Yet, most stochastic evaporation/evapotranspiration models relate evapotranspiration with net radiation, or other variables, such as minimum and maximum temperature, dew point temperature and wind speed, where these variables are obtained by conditioning them on the preceding day and the rainfall amount of the day considered (Lall et al. 1996; Srikanthan and McMahon 2001). However, through conditioning the different input variables (net radiation, temperature, ...) on the rainfall amount of the day considered, the correlation structure between evapotranspiration and precipitation is implicitly taken into account in these models. Alternative stochastic models of evapotranspiration make use of autoregressive models (often AR(3) models) (e.g. Alhassoun et al. 1997; Pandey et al. 2009) or autoregressive moving average (ARMA) models (e.g. Raghuwanshi and Wallender 1997), and do not account for precipitation. Furthermore, these time series models are used at monthly to yearly scale, making them inappropriate for use in rainfall-runoff models. In this research, we recognize the strong dependence of evapotranspiration on both net radiation and precipitation. However, given that air temperature is highly determined by net radiation, a high dependence of evapotranspiration on air temperature can be expected (as will be shown). As air temperature records are generally more accessible than net radiation time series, one could consider to

model evapotranspiration on the basis of temperature and precipitation.

Stochastic variables play a significant role in many hydrological processes (Salvadori and De Michele 2007). As these variables are usually not independent (e.g. storm intensity, magnitude and duration, or precipitation and evapotranspiration), it is important to be able to model this dependence in order to accurately estimate or analyse the risk involved in extremes or to derive time series of different variables that are in agreement. Such an analysis can be performed in a flexible and multivariate way by using copulas (Salvadori et al. 2007; Salvadori 2004; Salvadori and De Michele 2004). A copula (Sklar 1959; Nelsen 2006) is a multivariate function that describes the dependence structure between stochastic variables, independently of their marginal behaviour. As such, copulas do not suffer from the drawback that the marginal distribution functions have to belong to the same parametric family, and they permit the use of complex marginal distributions (Salvadori et al. 2007). Copulas have already proven their usefulness in hydrology. They have been employed in, for instance, the analysis of the dependence structure between storm characteristics (Vandenberghe et al. 2010a), in a statistical analysis of (extreme) rainfall events (Gräler et al. 2011; Vandenberghe et al. 2011; Kao and Govindaraju 2008) and in the development of stochastic rainfall models (Serinaldi 2009; Evin and Favre 2008; Salvadori and De Michele 2006; Vernieuwe et al. 2015). As copulas describe the dependence structure between stochastic variables, regardless of their marginal distributions, they are very useful for describing the dependences between evapotranspiration, rainfall characteristics, and other climatological variables such as net radiation or temperature.

The overall objective of this paper is to develop a stochastic evapotranspiration model that generates evapotranspiration time series that are in agreement with accompanying rainfall time series, such that it can be used in hydrological impact analysis. Based on stochastically generated rainfall and corresponding evapotranspiration time series, discharge series can be computed from a rainfall-runoff model. Hydrological impact analysis can then be based on the statistics of the extremes of the obtained discharge series. In order to be of use for this purpose, it is crucial that this new model preserves the statistical properties of the evapotranspiration time series and respect the dependence structure between precipitation and evapotranspiration. Furthermore, the model should be as simple as possible with respect to model input. Therefore, precipitation and daily temperature, two variables that are easy to measure or model, are selected as constraining variables for the evapotranspiration as was argued above. However, the model set-up should allow for replacing variables (e.g. net radiation instead of temperature) or

adding other variables (e.g. wind speed) that may influence evapotranspiration.

In order to develop a stochastic model, the dependence between the rainfall characteristics, temperature and evapotranspiration will first be described. To this end, time series of 72 years of data (precipitation, temperature and evapotranspiration) available for Uccle (Belgium) will be employed. As copulas model the dependence structure between different stochastic variables and have already proven their usefulness in hydrology, different copulas will be fitted, and their performance evaluated. Recently, vine copulas have been introduced (Bedford and Cooke 2001, 2002), i.e. multivariate parametric copulas built by the decomposition of the multivariate density into a product of bivariate copula densities. Given their properties, these copulas are preferred for describing the multivariate dependence structure between the aforementioned variables. Once the copulas are fitted, they will be used to generate time series of daily evapotranspiration values, given the recorded time series of rainfall records and daily temperature values, and their statistics will be compared with those of the observed evapotranspiration series to assess their modeling capacity.

The paper is structured as follows. Section 2 briefly introduces copulas and explains the copula-based simulation process. Section 3 introduces the observed time series that will be used for this study and presents the statistical dependence between the different variables considered. Section 4 describes the different models that are constructed, while Sect. 5 evaluates and discusses the simulations. Finally, Sect. 6 gives conclusions and recommendations for further investigations.

2 Copulas

A copula is a multivariate function that describes the dependence structure between random variables independent of their marginal behavior. As such, copulas do not suffer from the drawback that the marginal distribution functions have to belong to the same parametric family, and they permit the use of complex marginal distributions (Salvadori et al. 2007). The relation between bivariate distribution functions and bivariate copulas is given by the theorem of Sklar (Sklar 1959):

$$F_{12}(x_1, x_2) = C_{12}(F_1(x_1), F_2(x_2)) = C_{12}(u_1, u_2), \quad (1)$$

with F_{12} the joint cumulative distribution function of random variables X_1 and X_2 , C_{12} a bivariate copula, F_1 and F_2 two continuous marginal cumulative distribution functions of X_1 and X_2 , and $u_1 = F_1(x_1)$ and $u_2 = F_2(x_2)$. For more theoretical details, we refer to Sklar (1959) and Nelsen (2006).

Copulas have already proven their usefulness in hydrology, however, due to the complication in model construction for high-dimensional copula families, most research focuses on two variables. Only very limited applications can be found in the literature with multivariate analysis of rainfall (Gräler 2014; Gyasi-Agyei and Melching 2012; Zhang and Singh 2007; Kao and Govindaraju 2008; Salvadori and De Michele 2006; Grimaldi and Serinaldi 2006), floods (Zhang and Singh 2014; Xiong et al. 2014; Chen et al. 2012; Serinaldi and Grimaldi 2007; Genest et al. 2007; Salvadori and De Michele 2010) and droughts (Kao and Govindaraju 2010; Song and Singh 2010; Wong et al. 2010).

Recently, a flexible construction method for high-dimensional copulas, known as the vine copula (or pair-copula) construction (Kurowicka and Cooke 2007; Aas et al. 2009; Aas and Berg 2009; Haff et al. 2010), has been introduced and has shown a large potential for hydrological applications (e.g. Gräler et al. 2011; Vernieuwe et al. 2015). The advantage of the method is that it allows for constructing a multivariate copula based on the mixing of (conditional) bivariate copulas. In this paper, we restrict ourselves to vine copulas, although alternative multidimensional copulas could be used (e.g. the multivariate Gaussian copula). However, these often show less flexibility in describing the dependence structure between the variables considered.

The class of regular vine copulas is still very broad and embraces a large number of possible pair-copula decompositions (Aas et al. 2009). For example, there are 3, 24 and 240 different constructions for a three-, four- and five-dimensional vine copula respectively (Aas et al. 2009). There exist two special types of regular vine copulas: Canonical vine copulas (C-vine copulas) and D-vine copulas (Kurowicka and Cooke 2007). If all mutual dependences involve the same variable, the construction yields a C-vine copula. Figure 1 illustrates the principle of constructing three- and four-dimensional C-vine copulas. If all mutual dependences are considered one after the other, i.e. the first with the second, the second with the third, the third with the fourth, etc., the construction yields a D-vine copula. Only C-vine copulas were used in this study because of two reasons. Firstly, compared to D-vine copulas, C-vine copulas are easier to construct. Secondly, as temperature has the strongest relation with evapotranspiration (see the next paragraph), it is logical (Aas et al. 2009) to build a vine copula with temperature as the main variable.

In this paragraph, C-vine copulas are described in more detail. The construction of a four-dimensional C-vine copula is explained as follows. The pairwise dependences between the four variables U_1 , U_2 , U_3 and U_4 are captured by the bivariate copulas C_{12} , C_{13} and C_{14} , which is

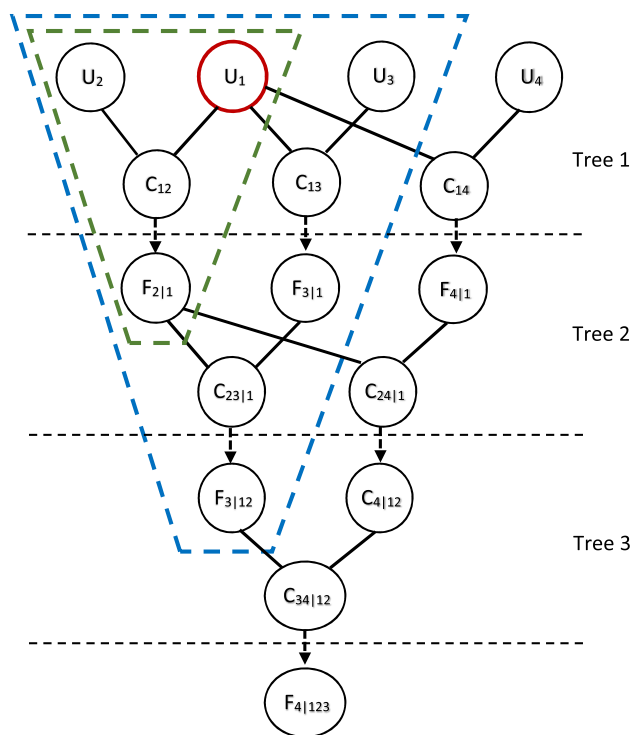


Fig. 1 Three- and four-dimensional C-vine copula. The three-dimensional C-vine copula is indicated within the blue dashed area

illustrated in the first tree of Fig. 1. These bivariate copulas can be conditioned on the variable U_1 through partial differentiation (Aas et al. 2009), resulting in the conditional cumulative distribution functions (CCDF) $F_{2|1}$, $F_{3|1}$ and $F_{4|1}$:

$$\begin{aligned} F_{2|1}(x_2|x_1) &= \frac{\partial}{\partial u_1} C_{12}(u_1, u_2); \\ F_{3|1}(x_3|x_1) &= \frac{\partial}{\partial u_1} C_{13}(u_1, u_3); \\ F_{4|1}(x_4|x_1) &= \frac{\partial}{\partial u_1} C_{14}(u_1, u_4). \end{aligned} \quad (2)$$

In the second tree, for all quadruples $(u_{1,i}, u_{2,i}, u_{3,i}, u_{4,i})$ the three conditional probabilities are then calculated ($i = 1, \dots, n$, with n the number of data points) and to these ‘conditioned observations’, which are again approximately uniformly distributed on $[0,1]$, two new bivariate copulas $C_{23|1}$ and $C_{24|1}$ are fitted. These copulas can also be conditioned by partial differentiation to obtain $F_{3|12}$ and $F_{4|12}$ in the third tree:

$$\begin{aligned} F_{3|12}(x_3|x_1, x_2) &= \frac{\partial C_{23|1}(F_{2|1}(x_2|x_1), F_{3|1}(x_3|x_1))}{\partial F_{2|1}(x_2|x_1)} \\ F_{4|12}(x_4|x_1, x_2) &= \frac{\partial C_{24|1}(F_{2|1}(x_2|x_1), F_{4|1}(x_4|x_1))}{\partial F_{2|1}(x_2|x_1)}. \end{aligned} \quad (3)$$

Finally, a bivariate copula $C_{34|12}$ is fitted, of which the partial derivative to $F_{3|12}$ can be computed to obtain the CCDF $F_{4|123}$.

Sampling values (u_1, u_2, u_3, u_4) out of a four-dimensional C-vine copula is straightforward and simple to implement. First, four random values (t_1, t_2, t_3, t_4) are independently drawn from a uniform distribution on $[0,1]$. These values are then used as probability levels of the CCDF Eqs. (4)–(7) on the basis of which (u_1, u_2, u_3, u_4) can be determined:

$$u_1 = t_1; \quad (4)$$

$$u_2 = F_{2|1}^{-1}(t_2|u_1); \quad (5)$$

$$u_3 = F_{3|1}^{-1}(F_{3|12}^{-1}(t_3|u_1, u_2)); \quad (6)$$

$$u_4 = F_{4|1}^{-1}(F_{4|12}^{-1}(F_{4|123}^{-1}(t_4|u_1, u_2, u_3))). \quad (7)$$

A simulation algorithm to draw random samples from a C-vine copula can be found in Aas et al. (2009).

3 Data set

At the Royal Meteorological Institute at Uccle near Brussels, Belgium, a 72-year time series (from 1931 to 2002) of daily reference evapotranspiration E is available which is derived from the Penman–Monteith method using on-site measured variables. Given the objective of the paper to develop a model for stochastic evapotranspiration generation based on rainfall and temperature data, time series of observed rainfall and mean daily temperature T are used as explanatory variables, while E is the response variable. It should be stated, however, that other variables, if available, could be used as well as explanatory variables. The precipitation data were extracted from the 105-year 10-minute rainfall time series (Demarée 2003), a data series that has been subjected to a large number of studies (Verhoest et al. 1997; Vaes et al. 2002; De Jongh et al. 2006; Ntegeka and Willems 2008; Vandenberghe et al. 2010b; Vanhaute et al. 2012; Pham et al. 2013; Willems 2013). These data have been reprocessed to daily total rainfall, further referred to as P and fraction of dry instances per day D , as both variables were believed to be correlated to daily evapotranspiration: a wet day with a negligible fraction of dry instances will show less evapotranspiration than another day having the same total rainfall amount but a high portion of dry periods (for instance due to a heavy intensity thunderstorm). In this study, we have restricted the explanatory variables to P , D and T (at day i) to predict E (at the same day i). As including the precipitation of the previous day ($i - 1$) in the analysis did not show any improvements in the model results (data not shown), it was

therefore not further considered as additional explanatory variable in the remainder of the paper. In order to avoid the seasonal effects in the data, the study investigated the dependence structures for each month separately.

Since the copulas are constructed based on the ranks of values, it is very critical to solve the problem of “ties” before fitting copulas to the data. The problem refers to the presence of events with identical values in the time series which has a large impact on the copula-fitting result (De Michele et al. 2007). In this study, “ties” commonly occurred during the period without rain or evapotranspiration. In order to remove “ties”, we used the method of adding “noise” as proposed by Vandenberghe et al. (2011). Values drawn uniformly at random from $[-0.001, 0.001]$ were added to the values of the variables. When $P = 0$ or $D = 0$ occurred, values drawn uniformly at random from $[0, 0.001]$ were added, whereas values drawn uniformly at random from $[-0.001, 0]$ were added when $D = 1$. Adding noise only results in negligible changes to the marginal distributions, yet resolving the problem of ties. More information about this problem can be found in Salvadori and De Michele (2006, 2007).

Table 1 presents the values of two common rank correlation coefficients that reflect the dependence between two variables, Kendall’s tau and Spearman’s rho, calculated for all variable combinations. The significance of the obtained values for Kendall’s tau and Spearman’s rho were tested as explained in Genest et al. (2007). All but three p values were smaller than 0.05, which indicates a dependence between the variables. From the table, it is clear that generally there is a strong correlation between E and T except for the months March, October and November. As can be expected, evapotranspiration is likely to be less

during wet days (and decreasing with increasing rainfall volumes) as such days are generally characterized by cloudy conditions and thus less energy (net radiation) that is available for evapotranspiration. The expected negative dependence between E and P is found for all months, except for the winter months December-January-February (DJF) for which a small positive correlation is obtained. Exactly the opposite was noticed for the relations between E and D : during spring, summer and autumn, evapotranspiration is positively correlated with the fraction of dry instances during the day, while during winter (DJF), small negative correlations were obtained.

4 Model construction

The stochastic evapotranspiration model that is developed in this paper employs a copula, such that, given a time series of rainfall and temperature data, a corresponding time series of evapotranspiration values can be generated by sampling the copula. As it is clear that E generally has a strong dependence with T , we propose to construct a C-vine copula having T as a core variable (as variable U_1 in Fig. 1). From Table 1, negative as well as positive dependences between T and the other explanatory variables P and D are observed. As these dependences should be respected within the model to be developed, a copula family should be selected that can describe positive as well as negative dependences. However, most common one-parameter bivariate copula families can only model positive dependence (Nelsen 2006), while only a few families allow for modeling negative dependences. In this paper, we first compare copulas determined by two different selection

Table 1 Values of Kendall’s tau and Spearman’s rho for all variables in each month

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
$EvsT$	0.516	0.514	0.328	0.405	0.526	0.455	0.467	0.433	0.371	0.246	0.191	0.505
	0.710	0.706	0.467	0.573	0.722	0.643	0.657	0.613	0.533	0.360	0.273	0.681
$EvsP$	0.130	0.040	-0.247	-0.358	-0.360	-0.378	-0.368	-0.357	-0.337	-0.239	-0.107	0.108
	0.193	0.058	-0.365	-0.521	-0.522	-0.545	-0.530	-0.522	-0.489	-0.353	-0.158	0.163
$EvsD$	-0.112	-0.024	0.257	0.381	0.379	0.406	0.401	0.380	0.357	0.258	0.122	-0.092
	-0.170	-0.039	0.376	0.548	0.546	0.577	0.570	0.549	0.511	0.380	0.178	-0.140
$TvsP$	0.268	0.192	0.037	-0.141	-0.172	-0.208	-0.224	-0.209	-0.139	-0.000	0.137	0.257
	0.396	0.286	0.055	-0.213	-0.258	-0.310	-0.329	-0.310	-0.206	0.001	0.203	0.381
$TvsD$	-0.264	-0.176	-0.028	0.154	0.186	0.230	0.248	0.236	0.148	0.002	-0.134	-0.250
	-0.392	-0.265	-0.042	0.231	0.277	0.343	0.365	0.352	0.218	0.001	-0.200	-0.370
$PvsD$	-0.800	-0.763	-0.747	-0.734	-0.706	-0.680	-0.687	-0.676	-0.706	-0.735	-0.802	-0.791
	-0.941	-0.916	-0.904	-0.898	-0.880	-0.861	-0.867	-0.858	-0.880	-0.897	-0.943	-0.936

Values of Kendall’s tau are indicated in bold

strategies. In a first strategy, we restrict ourselves to the Frank copula family as it allows for modeling the full range of dependences, and only requires one parameter that can easily be estimated. Furthermore, this copula family has frequently been used in hydrological applications (Pan et al. 2013). Also for the copula used in the bivariate copulas C_{PE} and C_{TE} (see further) we only use the Frank copula family in this first strategy.

In a second strategy, the copulas used within the C-vine copulas are selected on the basis of Akaike's information criterion (AIC) from six different copula families (the Gaussian, the t, the Clayton, the Gumbel, the Frank and the Joe family), hence allowing for a more flexible dependence structure. These vine copulas are further referred to as the 'optimal' vine copulas, although one must bear in mind that they do not necessarily represent a globally optimal fitted copula (Aas and Berg 2009; Nikoloulopoulos et al. 2012). In order to select the copula family to be used in the bivariate copulas C_{PE} and C_{TE} (see further), the Clarke test (Clarke 2007) was applied. This test was first employed by Belgorodski (2010) to calculate a goodness-of-fit score for selecting a copula family out of different families under consideration.

For both strategies, the bivariate copula parameters are estimated using the canonical maximum likelihood (CML) method that determines the parameter value θ that maximizes the rank-based likelihood function:

$$l(\theta) = \sum_{i=1}^n \ln \left[c \left(\frac{R_i^X}{n+1}, \frac{R_i^Y}{n+1} \right) \right], \quad (8)$$

with n the number of data points, R_i^X (respectively R_i^Y) the rank of the X -coordinate (respectively Y -coordinate) of the i -th data point and c the copula density function:

$$c = \frac{\partial^2 C}{\partial u \partial v}. \quad (9)$$

In order to randomly draw an evapotranspiration value that is conditioned on the explanatory variables T , P and D , a four-dimensional C-vine copula, referred to as V_{TPDE} (i.e. U_1 , U_2 , U_3 , and U_4 are derived from the marginal distributions of T , P , D , and E , respectively), is constructed. The alternative four-dimensional C-vine copula, V_{TDPE} was also assessed, but showed similar results to V_{TPDE} , and was therefore not further considered in this paper. To account for situations in which less data would be available, more simplified models are built as well. In case no sub-daily precipitation data are available from which D can be calculated, a three-dimensional C-vine copula, referred to as V_{TPE} , can be fitted that relates E to daily temperature T and daily precipitation P . Alternatively, if no temperature data would be available, E could be generated based on P and D data. In this case, a three-dimensional C-vine copula,

referred to as V_{PDE} , is constructed. Also the bivariate copulas C_{PE} is assessed, which could be used if only daily precipitation data would be available to relate E with.

The three-dimensional C-vine copula V_{TDE} (relating E to D and T) is not considered in this paper as it is unlikely to have time series of D , while daily precipitation data are not available. Nevertheless, this copula was tested and showed to behave similarly to V_{TPE} . Also, a bivariate copula C_{TE} will be included in the analysis in order to show the potential of this simple copula to simulate values of E based on daily temperatures. However, as will be shown, this copula does not allow for generating values of E that are consistent with occurring rainfall. Nevertheless, such a copula may be of use in applications where only time series of evapotranspiration are required or analysed regardless of precipitation. To show its performance, C_{TE} will be included even though simulations cannot be conditioned on precipitation.

Once the explanatory variables have been identified and the core variable is selected, a C-vine copula can be constructed. This will be demonstrated for the V_{TPDE} , since the other C-vine copulas follow the same method. In the first tree, U_1 , U_2 , U_3 and U_4 , derived from the marginal distributions of respectively T , P , D and E (see Fig. 1), were employed to select and fit the bivariate copulas C_{TP} , C_{TD} and C_{TE} . These bivariate copulas can be conditioned to the core variable (in this case T) through partial differentiation, resulting in the conditional cumulative distribution functions $F_{P|T}$, $F_{D|T}$ and $F_{E|T}$. In the second tree, the three conditional probabilities are then calculated for all data points. On these values, which are also uniformly distributed on $[0,1]$, the bivariate copulas $C_{PD|T}$ and $C_{PE|T}$ are selected and fitted. These copulas are then conditioned by partial differentiation to $F_{P|T}$ to obtain $F_{D|TP}$ and $F_{E|TP}$ in the third tree. Finally, a bivariate copula $C_{DE|TP}$ is selected and fitted, from which the partial derivative to $F_{D|TP}$ can be computed to obtain $F_{E|TPD}$.

5 Results and discussion

Table 2 illustrates which copulas were obtained in the 'optimal' bivariate and vine copulas as identified in the second selection strategy. This table shows that the Frank and Gaussian copulas are often selected. In order to find out whether the dependence present in the data is captured by the Frank C-vine copulas, i.e. the first strategy, the White goodness-of-fit test (Schepsmeier 2015) was applied to all these C-vine copulas. The testing and development of goodness-of-fit tests for vine copulas are still in its infancy. To our knowledge, only Schepsmeier (2015) investigated the performance of different goodness-of-fit tests for vine

Table 2 Copulas obtained in the ‘optimal’ bivariate and vine copulas as identified in the second selection strategy

Month	C_{TE}	C_{PE}	V_{TPE}			V_{PDE}			V_{TPDE}					
			C_{TP}	C_{TE}	$C_{EP T}$	C_{PD}	C_{PE}	$C_{DE P}$	C_{TP}	C_{TD}	C_{TE}	$C_{PD T}$	$C_{PE T}$	$C_{DE TP}$
Jan	F	F	F	F	F	F	C	t	F	F	F	F	F	F
Feb	F	F	F	F	F	F	C	t	F	F	F	F	F	F
Mar	t	C	C	t	F	F	F	F	C	F	t	F	F	F
Apr	F	F	F	G	F	F	F	F	F	Ga	G	F	F	F
May	F	F	F	Ga	F	F	F	F	F	Ga	Ga	F	F	F
Jun	F	F	F	Ga	F	F	F	F	F	C	Ga	F	F	t
Jul	F	F	F	Ga	F	F	F	C	F	C	Ga	F	F	C
Aug	F	F	F	G	F	F	F	F	F	F	G	F	F	C
Sep	F	F	Ga	Ga	F	F	Ga	F	Ga	t	Ga	F	F	F
Oct	F	F	Ga	Ga	F	F	F	F	Ga	C	Ga	F	F	F
Nov	t	t	F	t	Ga	F	t	t	F	F	t	F	Ga	t
Dec	F	F	Fa	t	Ga	F	F	F	F	F	t	F	Ga	F

F Frank copula, t t copula, Ga Gaussian copula, G Gumbel copula, C Clayton copula

copulas. He concluded that the White test performed very well. Application of this test to all Frank C-vine copulas determined in this research, yields p values ≥ 0.05 , indicating that the dependence structure of the data can be described by Frank copulas. As this goodness-of-fit test yields good results for the copulas obtained by the first strategy, and the copulas identified in the second strategy better fit the data in terms of AIC, the dependence in the data will also be preserved by these ‘optimal’ copulas.

In order to test the models’ performance, copula-based evapotranspiration simulations are performed and discussed. Given the historical observations of temperature and precipitation data, a copula-based simulation of values of E can be performed using one of the sampling algorithms Eqs. (4)–(7), i.e. only Eq. (6) is needed in case of a three-dimensional C-vine copula, and Eq. (7) for a four-dimensional C-vine copula. For example, the values of E simulated by the vine copula V_{TPDE} equals $F_{E|T}^{-1}(F_{E|TP}^{-1}(F_{E|TPD}^{-1}(t|u_t, u_p, u_d)))$, in which t is drawn from a uniform distribution on $[0, 1]$, u_t , u_p and u_d are obtained from the historical data of T , P and D through their empirical cumulative distribution functions (Vandenberghe et al. 2011). In this way, a simulation constitutes only a single realization of a stochastic process that is limited to the length of the time series (i.e. 72 year). Hence, the statistics of several simulations will show some variability. To account for these stochastic effects, the simulation is repeated 100 times for each copula. For each of the 100 simulations, the mutual dependences between the variables were assessed via Kendall’s tau. Figures 2 and 3 show box plots of the obtained values of Kendall’s tau, where values of E are estimated from observed data and their dependence on observed values of T or P is evaluated. These figures show that, generally, similar values of Kendall’s tau are obtained

for the Frank copulas and the ‘optimal’ copulas. Furthermore, it can be seen from these figures that, excluding P from the copulas, causes that the observed dependence between E and P is not captured. Therefore, the bivariate copulas C_{TE} are not suited to generate evapotranspiration time series to be used as forcing data in rainfall-runoff models since these data are not consistent with the precipitation data that are also used to force the model. The impact of using these data in order to model discharge, however, is outside the scope of this paper. Nevertheless, the bivariate copula C_{TE} is further included in the paper to assess its potential for stochastic generation of evapotranspiration time series for cases where the relation with precipitation is not required.

Figure 4 displays the comparisons between frequency distributions of observed and simulated evapotranspiration for the different months obtained by the Frank vine copulas V_{TPDE} . Similar figures showing minimal differences compared to those in Fig. 4 were found for the other copulas (i.e. V_{TPE} , V_{PDE} , C_{TE} and C_{PE}), and are therefore not shown. From the different plots, it can be seen that the frequency distribution of the observations of the reference evapotranspiration in Uccle (red line) is very similar to those obtained with the different copulas. However, due to the stochastic nature of the model and the fact that the simulated time series has a limited length, an individual distribution of simulated values may deviate from the observed distribution.

The simulations are further evaluated using the root mean square deviation (RMSD), given by:

$$\text{RMSD} = \sqrt{\frac{\sum_{i=1}^n (E_m(i) - E_o(i))^2}{n}}, \quad (10)$$

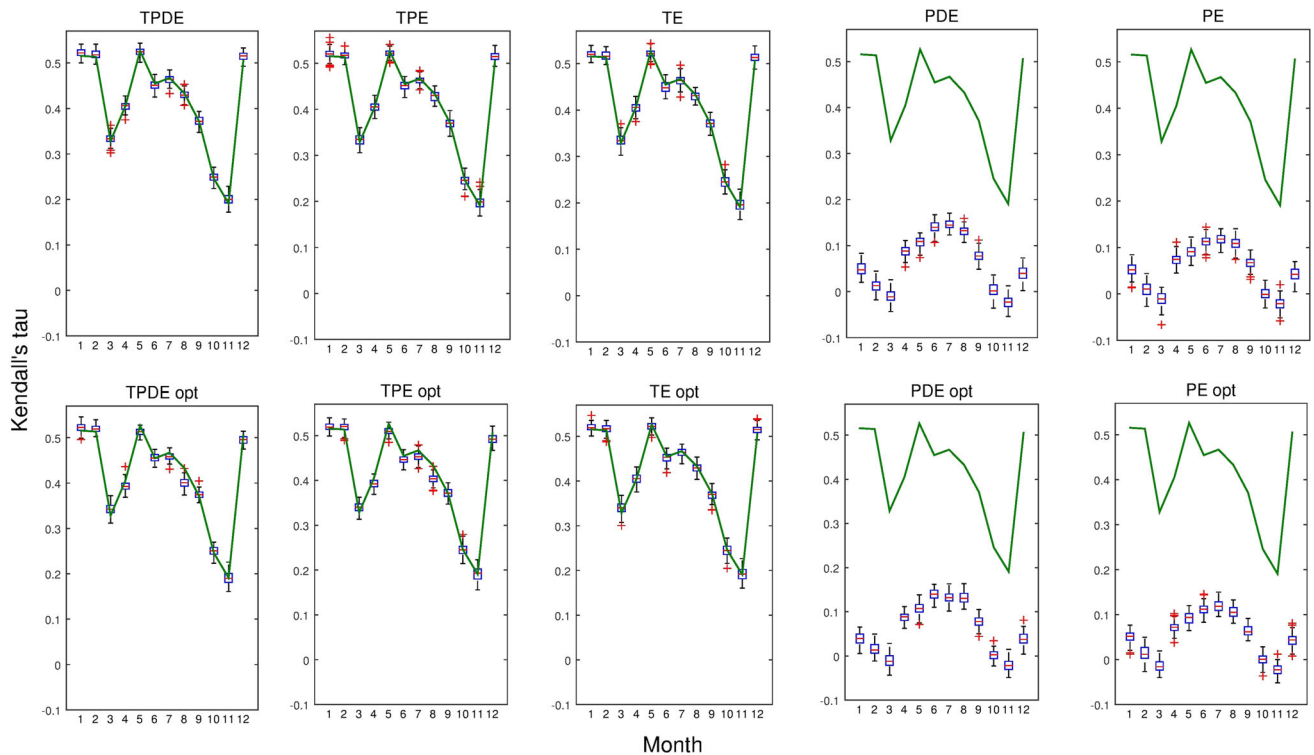


Fig. 2 Comparison between Kendall's tau for the relation between E and T of observed and simulated data for Frank copulas (top panel) and 'optimal' copulas (bottom panel): Uccle (green line), 100 simulated ensembles (box plot) for V_{TPDE} , V_{TPE} , C_{TE} , V_{PDE} and C_{PE}

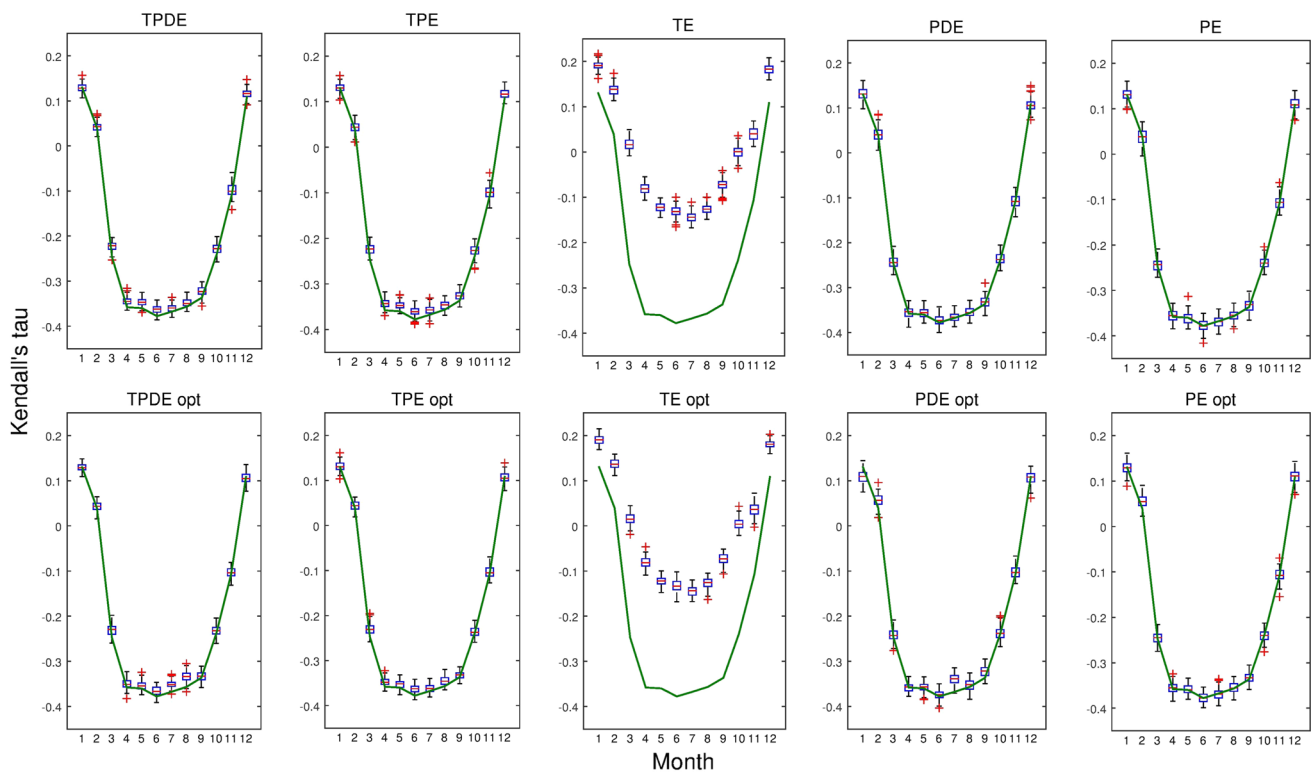


Fig. 3 Comparison between Kendall's tau for the relation between E and P of observed and simulated data for Frank copulas (top panel) and 'optimal' copulas (bottom panel): Uccle (green line), 100 simulated ensembles (box plot) for V_{TPDE} , V_{TPE} , C_{TE} , V_{PDE} and C_{PE}

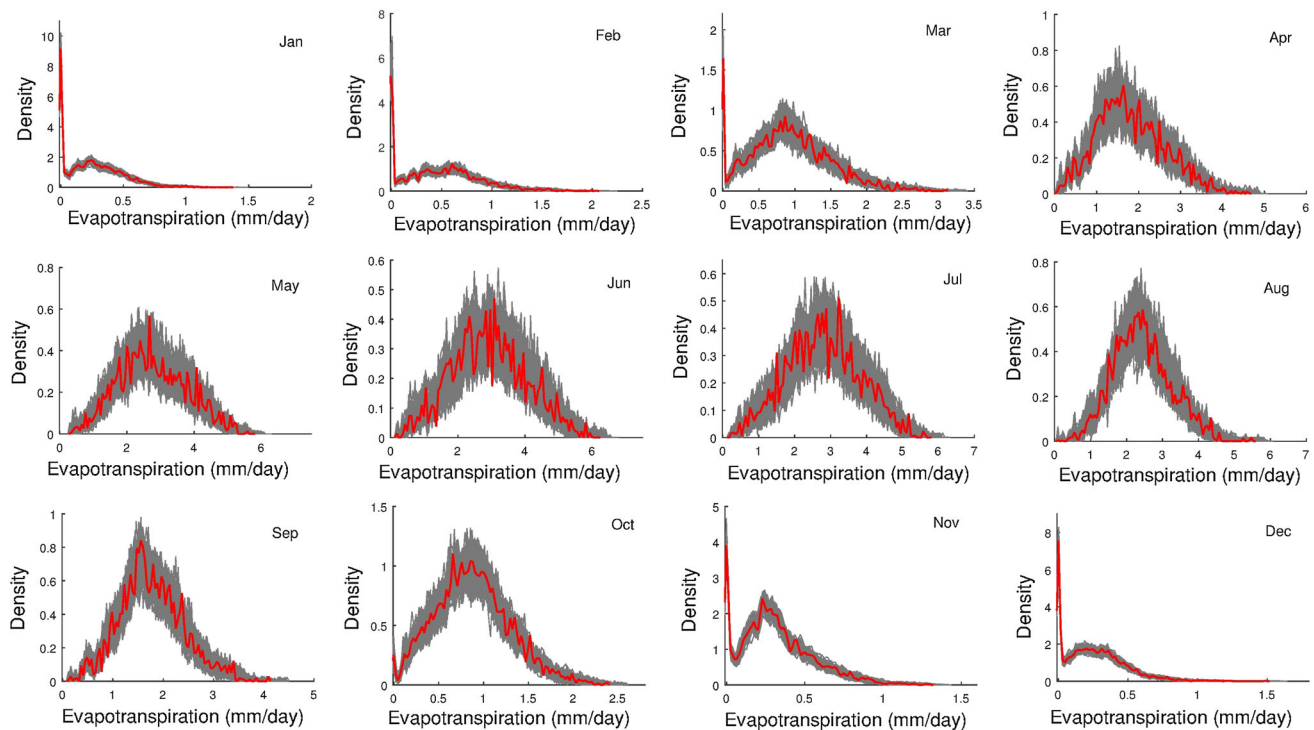


Fig. 4 Comparison between the frequency distributions of E of observed and simulated data: Uccle (red), 100 ensembles simulated using the Frank vine copulas V_{TPDE} (grey)

where $E_m(i)$ and $E_o(i)$ are respectively the modeled and observed evapotranspiration value at instant i and n is the number of values considered to calculate the RMSD upon.

The results of the 100 simulations using each copula are summarized as box plots in Fig. 5. For all copulas, the largest deviations occur during the period from April to September. However, during these months (spring to autumn), larger evapotranspiration values are found and deviations compared to the observed time series should be interpreted relative to the mean E during the month considered. Figure 6 shows these relative deviations as relative RMSD (RRMSD) values that equal the RMSD divided by the average value of E for the month considered. As can be seen during winter months, the deviations are of the same order or larger than the average evapotranspiration, while in summer months, these deviations reduce to less than 40 % of the average value (in case of V_{TPDE}). It should be stated that the RMSD cannot be interpreted as an error as the models do not try to predict the observations. The RMSD merely formulates how a model realization deviates from the observations. Given that both the observations and the model realization result from stochastic processes, it cannot be expected that they are exactly the same. However, smaller values of RMSD (or RRMSD) correspond to models that behave more similar to the observations than model realizations with higher values of both statistics. The

RMSD can thus be used to rank the copulas. It can be seen that the vine copulas V_{TPDE} and V_{TPE} are ranked above the other copulas, yet, there is no major difference between V_{TPDE} and V_{TPE} , which indicates that adding D as explanatory variable does not improve the performance. Figure 5 shows that all copulas that include temperature as an explanatory variable perform better than those that are only based on precipitation, and that, including less explanatory variables enlarges the deviations of the individual model realizations with respect to the observed time series. The worst copula developed is C_{PE} . Including fraction drought to this copula (resulting in the vine copula V_{PDE}) improves the performance, though the vine copula V_{PDE} is still worse than any other copula that uses daily temperature as input. Furthermore, one can see that the performance of models using the Frank copulas and the ‘optimal’ copulas cannot be distinguished visually. On this basis, and given the result of the goodness-of-fit tests on the Frank vine copulas, one can conclude that for this case study, there is no major improvement of working with more flexible vine copulas. For reasons of ease and simplicity, we opted to exclude the ‘optimal’ bivariate and vine copulas for the remainder of the paper.

Figures 7 and 8 display spaghetti-plots of the 100 simulations using the copulas V_{TPDE} , V_{TPE} , C_{TE} , V_{PDE} and C_{PE} for a simulation of 5 years (1998–2002) of evapotranspiration

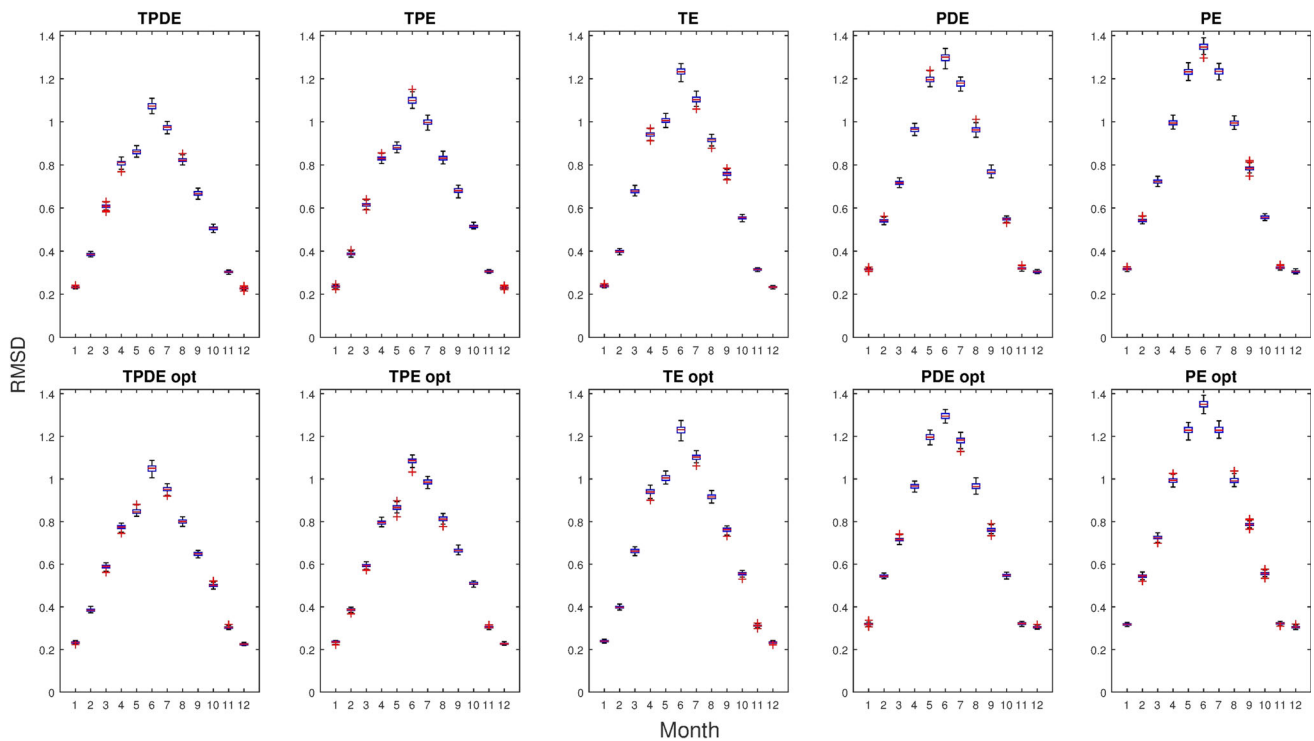


Fig. 5 Box plots of RMSD of E simulated by V_{TPDE} , V_{TPE} , C_{TE} , V_{PDE} and C_{PE} for the Frank copulas (top panel) and 'optimal' copulas (bottom panel)

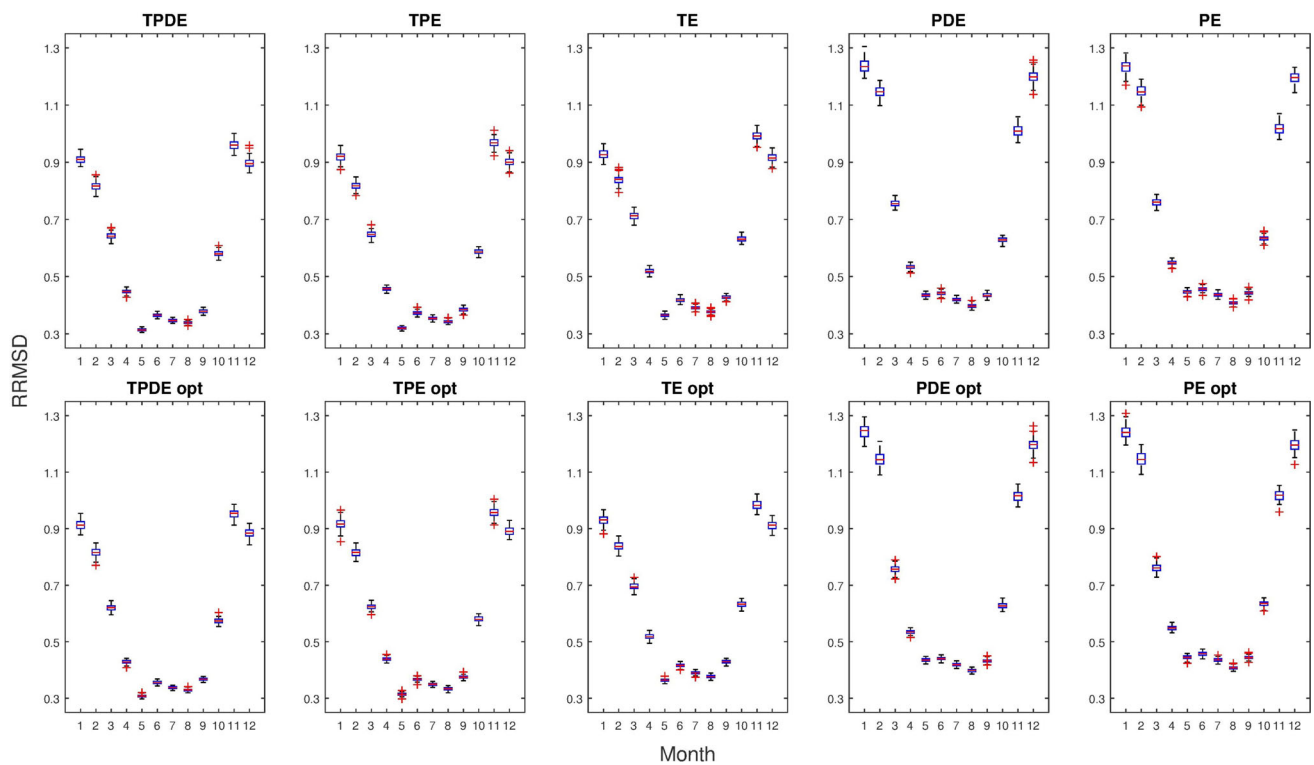


Fig. 6 Box plots of RRMSD of E simulated by V_{TPDE} , V_{TPE} , C_{TE} , V_{PDE} and C_{PE} for the Frank copulas (top panel) and 'optimal' copulas (bottom panel)

during the months of January (characterized by the smallest RMSD) and June (having the largest RMSD), respectively. Also included in these figures is the observed evapotranspiration time series (black line). It is clear from these figures that the observations always fall within the ensemble range and that the average of the ensembles is close to the observed time series, except for the copulas C_{PE} and V_{PDE} . The latter copulas are not able to estimate trends in E : periods characterized by low or high values of evapotranspiration are not captured by the copula, and for the month of January (but also other winter months—data not shown), a too low temporal variability is generated. Comparing the different figures, it is clear that smaller ensemble ranges are obtained for the copulas involving T (with V_{TPDE} the copula with the smallest range), while simulations using the copulas V_{PDE} and C_{PE} show large ensemble ranges that hardly follow the trend in the observed time series. This behavior reveals that, at all times, the latter copulas simulate values that may be very different from the observations. The reason for this improper behavior should be sought in the fact that the dependence between E and precipitation-related variables (P and D) is too small to constrain the evapotranspiration-generating process.

Conclusions cannot solely be made based on the ensemble width of the spaghetti-plots and the temporal behavior of the ensemble mean as both do not fully allow for evaluating the model behavior. To get a better insight, we randomly highlighted one realization (in cyan) to show how its temporal variability compares to that of the observed time series. Based on a visual appreciation of these figures (and this can be confirmed from the RMSD results discussed above), the simulations using the vine copula V_{TPDE} show a similar temporal behavior as the observations, while decreasing the number of explanatory variables causes rapid temporal changes in modeled evapotranspiration. In this respect, the copulas V_{PDE} and C_{PE} behave the worst.

To further assess the copulas, the mean daily evaporation for each month was calculated for each ensemble member and compared to the mean daily evaporation at Uccle. Figure 9, displaying these results, shows that all copulas are capable of preserving the long-term mean monthly mean well (i.e. calculated from 72 year of data), and that very small differences are found between the ensemble members. In order to assess the variability in the modeled series, the standard deviation of the daily evapotranspiration of each ensemble member, calculated for the different copulas, was compared to that of the observations (cfr. Fig. 10). As can be seen from Fig. 10, all modeled

series show similar standard deviations at the daily level. However, when the standard deviations of the monthly total evaporation are compared to those of the observations, we find that all modeled series underestimate this monthly variability (cfr. Fig. 11). For the copulas involving T (i.e. V_{TPDE} , V_{TPE} and C_{TE}), these underestimations are fairly small, while for the copulas C_{PE} and V_{PDE} , the variability is much too small. The modeled series using the latter copulas insufficiently capture the annual variability of the evapotranspiration, as this variability is insufficiently reflected in the precipitation data. Daily temperature allows for introducing this interannual variability, though larger variabilities are still needed, signifying that the information content in daily temperature may not be sufficient. Other data that give more information on the temperature during the period of evapotranspiration (i.e. daytime), such as maximum temperature or mean daytime temperature, might lead to better models. Further extending the copulas with variables that directly influence the evaporation process, such as net radiation and wind speed, may further improve the modeling. However, it might be difficult to obtain such data sets from observations or from stochastic models.

Figure 12 presents the comparison between observed E and the ensemble mean for the different copulas considered for two years (i.e. 1931–1932). The results from the copulas involving T seem to be very similar and close to the reference evapotranspiration. From this figure, it is clear that the ensemble means for the copulas C_{PE} and V_{PDE} show too small a variability in the winter months. During the other months, these ensemble means show a larger variability, though they are not very consistent with the observations (e.g. during the period of high evapotranspiration in April 1932, the different ensembles do not consistently follow these higher observed values, while for the copulas involving T , all ensemble members simulate larger values of E).

Finally, the different copulas are evaluated by comparing the ensemble average total annual evapotranspiration to the annual reference evapotranspiration observed at Uccle (see Fig. 13). Taking into account the smoothing effect when averaging, it can be concluded that all copulas involving T seem to be able to preserve the annual evapotranspiration well. Again, the copula that uses most information for constraining the evapotranspiration simulations, i.e. V_{TPDE} , remains closest to the reference data, followed by the simulations obtained by V_{TPE} and C_{TE} . The copulas C_{PE} and V_{PDE} are not able to mimic the yearly variability in total evapotranspiration.

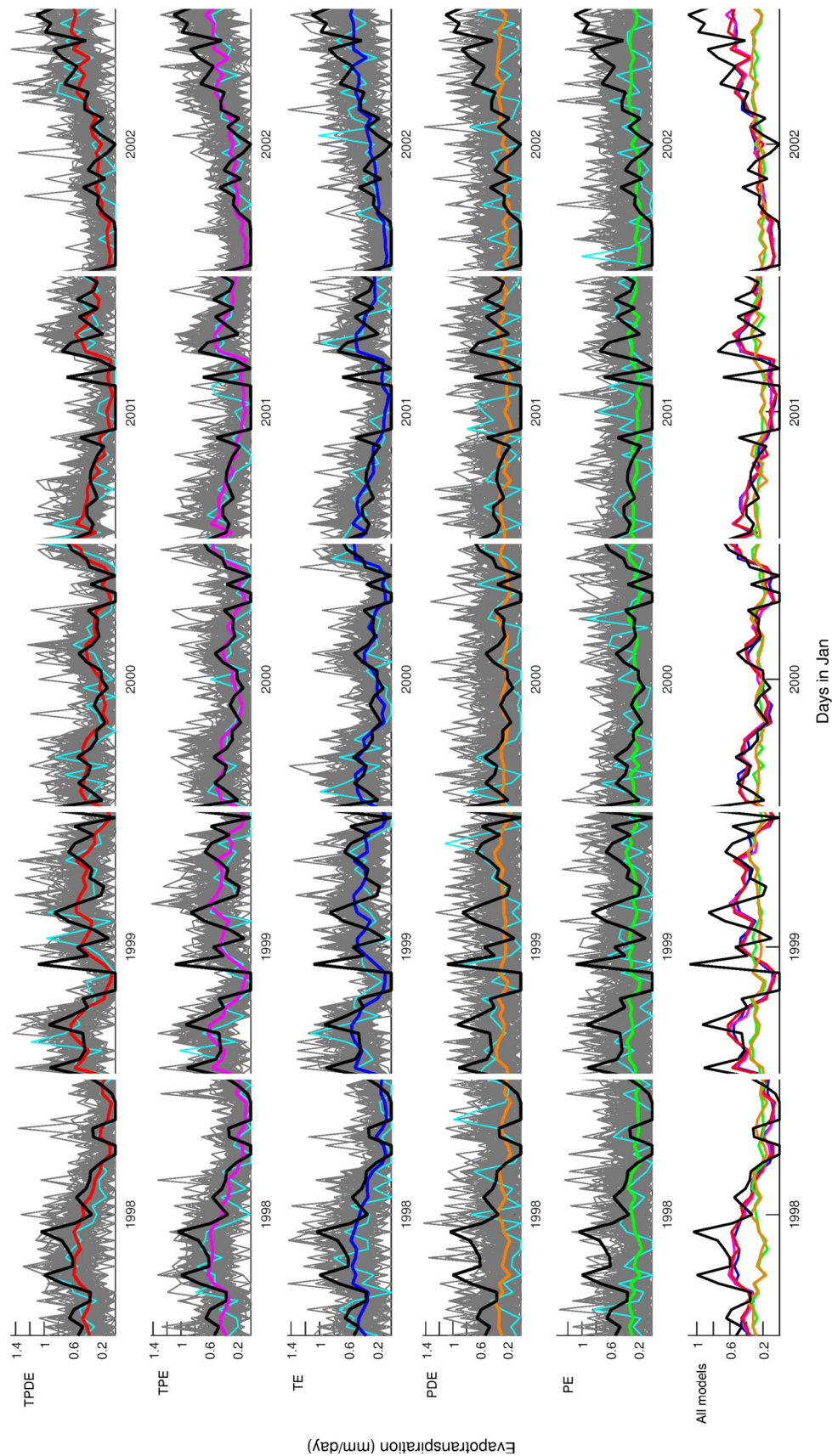


Fig. 7 Comparison between the observed and simulated time series of *E*: Uccle (*black*), 100 ensembles simulated using the different Frank copulas (*gray*), a random simulation ensemble for each copula (*cyan*), mean of 100 V_{TPE^-} (*red*), 100 V_{TPE^-} (*magenta*), 100 C_{TE^-} (*blue*), 100 V_{PDE^-} (*orange*) and C_{PDE^-} (*green*) simulated ensembles for the month of January during the last 5 years (1998–2002)

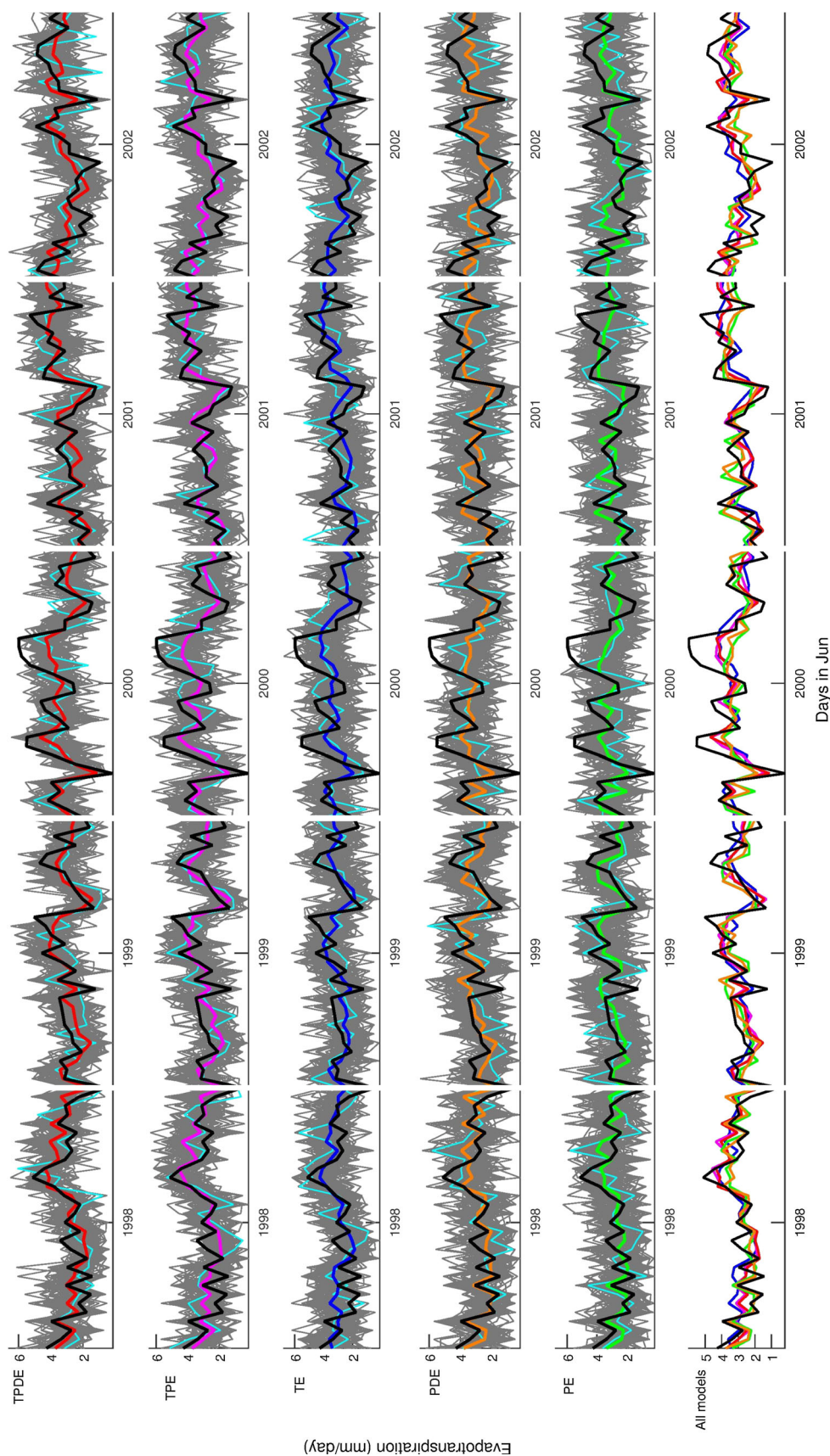


Fig. 8 Comparison between the observed and simulated time series of E : Uccle (black), 100 ensembles simulated using the different Frank copulas ($gray$), a random simulation ensemble for each copula (cyan), mean of V_{TPDE} (red), 100 V_{TPDE} (magenta), 100 C_{TE} (blue), 100 V_{PDE} (orange) and C_{PE} (green) simulated ensembles for the month of June during the last 5 years (1998–2002)

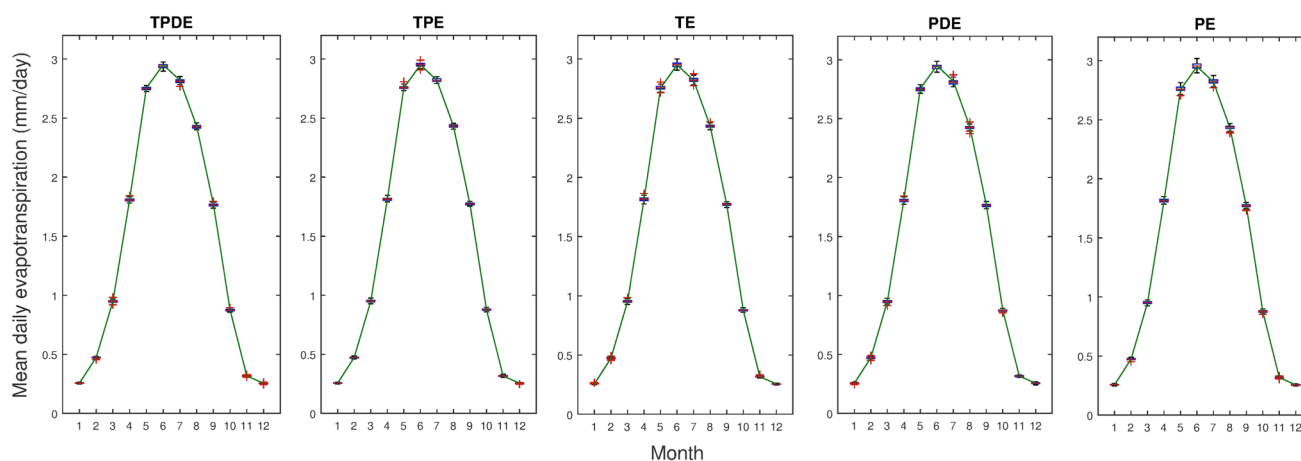


Fig. 9 Box plots of the daily mean evapotranspiration for the different months. The *green line* represents the average daily evapotranspiration observed at Uccle, Belgium

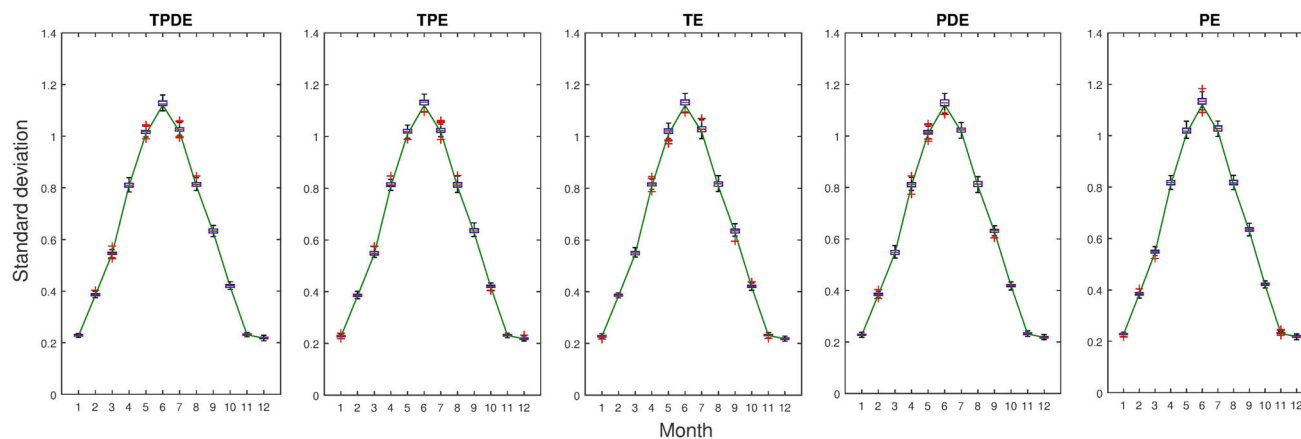


Fig. 10 Box plots of the standard deviation of daily evapotranspiration for the different months. The *green line* corresponds to the observations at Uccle, Belgium

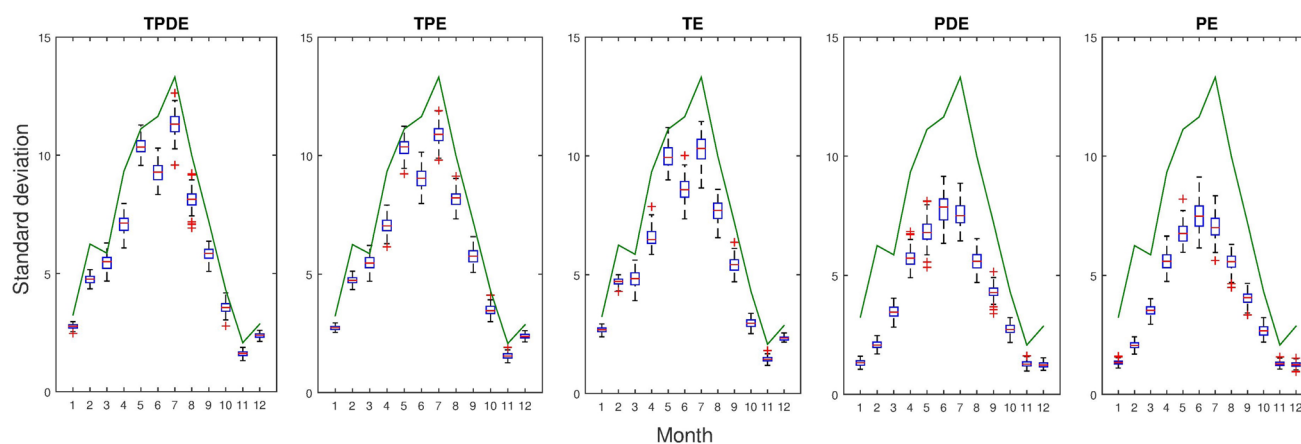


Fig. 11 Box plots of the standard deviation of total monthly evapotranspiration for the different months. The *green line* corresponds to the observations at Uccle, Belgium

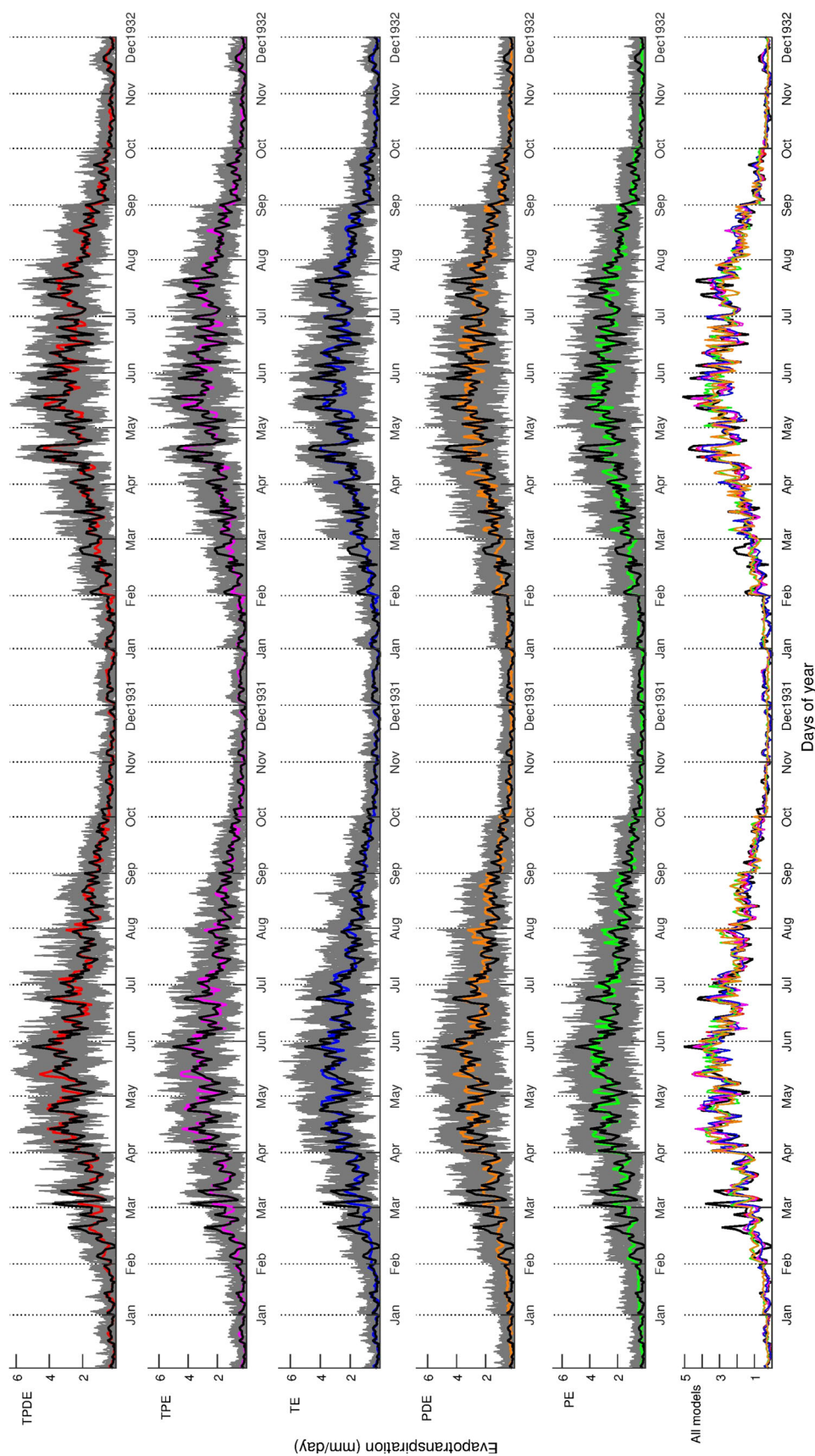


Fig. 12 Comparison between the observed and simulated time series of E using Frank copulas: Uccle (black), mean of 100 V_{TPDE} - (red), 100 V_{TPE} - (magenta), 100 C_{TE} - (blue), 100 V_{PDE} - (orange) and 100 C_{PE} - (green) simulated ensembles during 1931–1932

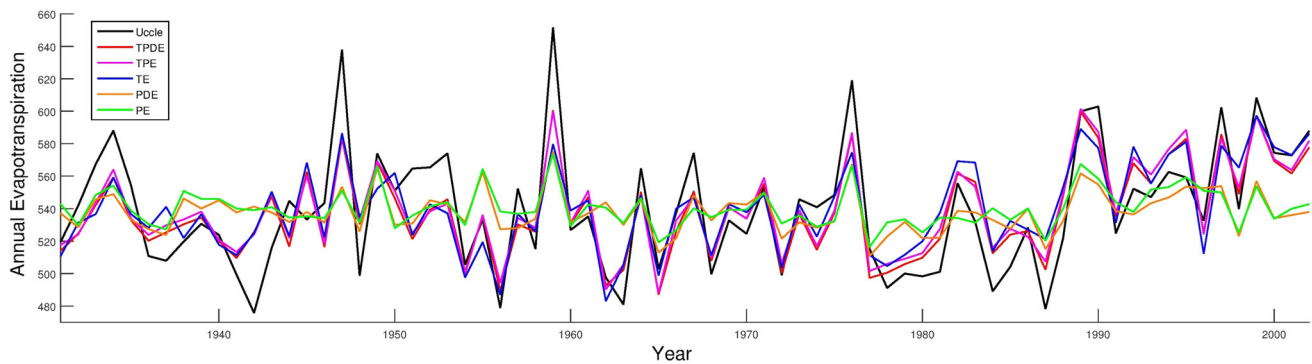


Fig. 13 Comparison between the total annual reference evapotranspiration of Uccle (black) and the average annual evapotranspiration of simulated data using Frank copulas V_{TPDE} (red), V_{TPE} (magenta), C_{TE} (blue) V_{PDE} (orange) and C_{PE} (green)

6 Conclusion and recommendation

Along with precipitation, evapotranspiration is a very important component in the water balance and therefore has a large impact on the catchment discharge. In order to assess extreme statistics of the discharge for water management planning and decision making, extremely long time series of precipitation and evapotranspiration may be required as inputs to hydrological models. One can make use of stochastic point process rainfall models to obtain the rainfall time series, however, a stochastic evapotranspiration model that provides evapotranspiration time series that are not in conflict with the rainfall time series, has not been developed yet. In this paper, different models were developed in which besides precipitation data also temperature data were used to constrain the evapotranspiration values.

Based on a record of 72 year (1931–2002) daily temperature T , precipitation P , dry fraction D and reference evapotranspiration E for Uccle in Belgium, several copulas were fitted. A four-dimensional C-vine copula, V_{TPDE} , two three-dimensional C-vine copulas, V_{TPE} and V_{PDE} , and two bivariate copulas C_{TE} and C_{PE} were considered. Given time series of T , P and D , V_{TPDE} provides stochastic values of E that are constrained by the T , P and D values, while for V_{TPE} , E is generated conditional to T and P . For the V_{PDE} , E is constrained by P and D . For both bivariate copulas C_{TE} and C_{PE} , evapotranspiration is generated conditional to respectively T and P . Regarding the choice of copula families to be used, two strategies were followed. In a first strategy, only Frank copulas were selected. In a second strategy, optimal copulas were selected from six different copula families on the basis of the AIC in order to obtain more flexible dependence models. Results showed that the dependence structure of the data is supported by models originating from both strategies. Also, no visual difference in terms of RMSD and RRMSD could be observed between both strategies which led to the decision to only include the

simpler Frank copulas for the remainder of the paper. From the analyses, it was furthermore found that all copulas involving T (i.e. V_{TPDE} , V_{TPE} and C_{TE}) provide acceptable simulations, where including more explanatory variables provide better models. Still, as no major difference in performance between simulations using V_{TPDE} and V_{TPE} was observed, the benefit of adding D to the copulas can be questioned. The copulas involving P (i.e. V_{PDE} and C_{PE}) showed not to be able to preserve certain trends (periods of high or low evapotranspiration). However, the bivariate copula C_{TE} cannot be used for applications where a simultaneous use of both evapotranspiration and precipitation time series are required, as it cannot guarantee a correct dependence between the modelled evapotranspiration and the precipitation time series. Only in cases where only evapotranspiration time series are required and no precipitation data are available, modeling E based on the bivariate copula C_{TE} through conditioning it on observed temperature values is a worthy alternative.

From this study, we may thus conclude that, in order to generate long-term evapotranspiration time series that correctly accompany stochastic rainfall series, one should rely on both a stochastically generated rainfall series and a temperature generator. Based on time series from these models, T and P (and D) data can be derived as input to the T -based copulas V_{TPDE} or V_{TPE} (the latter in case the rainfall generator does not provide subdaily data). However, the copulas developed can still be extended with other data that show correlations with the evapotranspiration (e.g. maximum daily temperature, net radiation, wind speed,...). Through adding more explanatory variables, copulas can be obtained that even better preserve the evapotranspiration statistics.

Acknowledgments The authors gratefully acknowledge the Vietnamese Government Scholarship (VGS) and the project G.0013.11N of the Research Foundation Flanders (FWO) for their partial financial support for this work. The historical Uccle series were provided by the Royal Meteorological Institute of Belgium.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Aas K, Berg D (2009) Models for construction of multivariate dependence—a comparison study. *Eur J Financ* 15(7–8):639–659
- Aas K, Czado C, Frigessi A, Bakken H (2009) Pair-copula constructions of multiple dependence. *Insurance* 44(2):182–198
- Alhassoun S, Sendil U, Al-Othman AA, Negm AM (1997) Stochastic generation of annual and monthly evaporation in Saudi Arabia. *Can Water Resour J* 22(2):141–154
- Arora VK (2002) The use of the aridity index to assess climate effect on annual rainfall runoff. *J Hydrol* 265:164–177
- Bedford R, Cooke RM (2001) Probability density decomposition for conditionally dependent random variables modeled by vines. *Ann Math Artif Intell* 32:245–268
- Bedford R, Cooke RM (2002) Vines—a new graphical model for dependent random variables. *Ann Stat* 30:1031–1068
- Belgorodski N (2010) Selecting pair-copula families for regular vines with application to the multivariate analysis of European stock market indices. Master's thesis, Technische Universitaet Muenchen
- Boughton W, Droop O (2003) Continuous simulation for design flood estimation—a review. *Environ Model Softw* 18(4):309–318
- Chen L, Singh VP, Guo SL, Hao ZC, Lli TY (2012) Flood coincidence risk analysis using multivariate copula functions. *J Hydrol Eng* 17(6):742–755
- Clarke KA (2007) A simple distribution-free test for nonnested model selection. *Political Anal* 15:347–363
- De Jongh ILM, Verhoest NEC, De Troch FP (2006) Analysis of a 105-year time series of precipitation observed at Uccle, Belgium. *Int J Climatol* 26(14):2023–2039
- De Michele C, Salvadori G, Passoni G, Vezzoli R (2007) A multivariate model of sea storms using copulas. *Coast Eng* 54(10):734–751
- Demarée G (2003) The centennial recording raingauge of the Uccle Plateau: its history, its data and its applications. *La Houille Blanche - Revue Internationale de l'Eau* 4:95–102
- Evin G, Favre AC (2008) A new rainfall model based on the Neymann-Scott process using cubic copulas. *Water Resour Res* 44(W03):433
- Genest C, Favre AC, Beliveau J, Jacques C (2007) Metaelliptical copulas and their use in frequency analysis of multivariate hydrological data. *Water Resour Res* 43(W09):401
- Gerrits AMJ, Savenije HHG, Veling EJM, Pfister L (2009) Analytical derivation of the Budyko curve based on rainfall characteristics and a simple evaporation model. *Water Resour Res* 45(W04):403
- Gräler B (2014) Modelling skewed spatial random fields through the spatial vine copula. *Spatial Statistics* 10:87–102
- Gräler B, van den Berg MJ, Vandenberghe S, Petroselli A, Grimaldi S, De Baets B, Verhoest NEC (2011) Multivariate return periods in hydrology: a critical and practical review focusing on synthetic design hydrograph estimation. *Hydrol Earth Syst Sci* 17:1281–1296
- Grimaldi S, Serinaldi F (2006) Design hyetograph analysis with 3-copula function. *Hydrol Sci J* 51(2):223–238
- Gyasi-Agyei Y, Melching CS (2012) Modelling the dependence and internal structure of storm events for continuous rainfall simulation. *J Hydrol* 464:249–261
- Haff IH, Aas K, Frigessi A (2010) On the simplified pair-copula construction—simply useful or too simplistic? *J Multivar Anal* 101(5):1296–1310
- Jones JW, Colwick RE, Threadgill ED (1972) A simulated environmental model of temperature, evaporation, rainfall and soil moisture. *Trans ASAE* 15:366–372
- Kao SC, Govindaraju RS (2008) Trivariate statistical analysis of extreme rainfall events via the Plackett family of copulas. *Water Resour Res* 44(W02):415
- Kao SC, Govindaraju RS (2010) A copula-based joint deficit index for droughts. *J Hydrol* 380:121–134
- Kurowicka E, Cooke RM (2007) Sampling algorithms for generating joint uniform distributions using the vine-copula method. *Comput Stat Data Anal* 51(6):2889–2906
- Lall U, Rajagopalan B, Tarboton DG (1996) A nonparametric wet/dry spell model for resampling daily precipitation. *Water Resour Res* 32(9):2803–2823
- Nelsen RB (2006) An introduction to copulas. Springer, New York
- Nikoloulopoulos AK, Joe H, Li H (2012) Vine copulas with asymmetric tail dependence and applications to financial return data. *Comput Stat Data Anal* 56:3659–3673
- Ntegeka V, Willems P (2008) Trends and multidecadal oscillations in rainfall extremes, based on a more than 100-year time series of 10 min rainfall intensities at Uccle, Belgium. *Water Resour Res* 44(W07):402
- Pan M, Yuan X, Wood EF (2013) A probabilistic view on drought recovery. *Geophys Res Lett* 40:3637–3642
- Pandey P, Pandey V, Singh R, Bhakar SR (2009) Stochastic modelling of actual Black Gram evapotranspiration. *J Water Resour Prot* 1(6):448–455
- Pham MT, Vanhaute WJ, Vandenberghe S, De Baets B, Verhoest NEC (2013) An assessment of the ability of Bartlett-Lewis type of rainfall models to reproduce drought statistics. *Hydrol Earth Syst Sci* 17:5167–5183
- Raghuwanshi NS, Wallender WW (1997) Field-measured evapotranspiration as a stochastic process. *Agric Water Manag* 32(2):111–129
- Salvadori G (2004) Bivariate return periods via 2-copulas. *Stat Methodol* 1:129–144
- Salvadori G, De Michele C (2004) Frequency analysis via copulas: theoretical aspects and applications to hydrological events. *Water Resour Res* 40(W12):511
- Salvadori G, De Michele C (2006) Statistical characterization of temporal structure of storms. *Adv Water Resour* 29(6):827–842
- Salvadori G, De Michele C (2007) On the use of copulas in hydrology. *J Hydrol Eng* 12(4):369–380
- Salvadori G, De Michele C (2010) Multivariate multiparameter extreme value models and return periods: a copula approach. *Water Resour Res* 46(W10):501
- Salvadori G, De Michele C, Kottegoda NT, Rosso R (2007) Extremes in nature: an approach using copulas. Springer, New York
- Schepsmeier U (2015) Efficient information based goodness-of-fit tests for vine copula models with fixed margins: a comprehensive review. *J Multivar Anal* 138:34–52
- Serinaldi F (2009) A multisite daily rainfall generator driven by bivariate copula-based mixed distributions. *J Geophys Res Atmos* 114:D10103
- Serinaldi F, Grimaldi S (2007) Fully nested 3-copula: procedure and application on hydrological data. *J Hydrol Eng* 12(4):420–430
- Sklar A (1959) Fonctions de repartition à n dimensions et leurs marges. *Publications de l'Institut de Statistique de l'Université de Paris* 8:229–231
- Song S, Singh VP (2010) Meta-elliptical copulas for drought frequency analysis of periodic hydrologic data. *Stoch Env Res Risk Assess* 24:245–444

- Srikanthan R, McMahon TA (2001) Stochastic generation of annual, monthly and daily climate data: a review. *Hydrol Earth Syst Sci* 5(4):653–670
- Vaes G, Willems P, Berlamont J (2002) 100 years of Belgian rainfall: are there trends? *Water Sci Technol* 45(2):55–61
- Vandenberghe S, Verhoest NEC, Buyse E, De Baets B (2010a) A stochastic design rainfall generator based on copulas and mass curves. *Hydrol Earth Syst Sci* 14:2429–2442
- Vandenberghe S, Verhoest NEC, De Baets B (2010b) Fitting bivariate copulas to the dependence structure between storm characteristics: a detailed analysis based on 105 year 10 min rainfall. *Water Resour Res* 46(W01):512
- Vandenberghe S, Verhoest NEC, Onof C, De Baets B (2011) A comparative copula-based bivariate frequency analysis of observed and simulated storm events: a case study on Bartlett-Lewis modeled rainfall. *Water Resour Res* 47(W07):529
- Vanhaute WJ, Vandenberghe S, Scheerlinck K, De Baets B, Verhoest NEC (2012) Calibration of the modified Bartlett-Lewis model using global optimization techniques and alternative objective functions. *Hydrol Earth Syst Sci* 16:873–891
- Verhoest N, Troch PA, De Troch FP (1997) On the applicability of Bartlett-Lewis rectangular pulses models in the modeling of design storms at a point. *J Hydrol* 202:108–120
- Verhoest NEC, Vandenberghe S, Cabus P, Onof C, Meca-Figueras T, Jameleddine S (2010) Are stochastic point rainfall models able to preserve extreme flood statistics? *Hydrol Process* 24:3439–3445
- Vernieuwe H, Vandenberghe S, De Baets B, Verhoest NEC (2015) A continuous rainfall model based on vine copulas. *Hydrol Earth Syst Sci* 19(6):2685–2699
- Wheater HS (2002) Progress in and prospects for fluvial flood modeling. *Philos Trans R Soc Lond A* 360(1796):1409–1431
- Willems P (2013) Adjustment of extreme rainfall statistics accounting for multidecadal climate oscillations. *J Hydrol* 490:126–133
- Willems P (2014) Parsimonious rainfall-runoff model construction supported by time series processing and validation of hydrological extremes Part 1: step-wise model-structure identification and calibration approach. *J Hydrol* 510:578–590
- Wong G, Lambert MG, Leonard M, Metcalfe AV (2010) Drought analysis using trivariate copulas conditional on climatic states. *J Hydrol Eng* 15(2):129–141
- Xiong L, Yu K, Gottschalk L (2014) Estimation of the distribution of annual runoff from climatic variables using copulas. *Water Resour Res* 50:7134–7152
- Zhang L, Singh VP (2007) Gumbel-Hougaard copula for trivariate rainfall frequency analysis. *J Hydrol Eng* 12(4):409–419
- Zhang L, Singh VP (2014) Trivariate flood frequency analysis using discharge time series with possible different lengths: Cuyahoga river case study. *J Hydrol Eng* 19(10):05014012